THE EXISTENCE OF STATIONARY MEASURES

FOR CERTAIN MARKOV PROCESSES

By

T. E. Harris

P-728

31 August 1955

SUMMARY

The idea of a stationary distribution for a Markov
process can be extended to include distribution or measures
which are infinite. It is shown that a certain type of recur-
rence condition implies the existence of a possibly infinite
stationary measure. Applications are discussed.

ACKNOWLEDGMENT

THE EXISTENCE OF STATIONARY MEASURES FOR CERTAIN
MARKOV PROCESSES

1. Introduction

We consider a Markov process x_n, n = 0,1,.... The random variables x_n belong to an abstract set X in which a Borel field B is defined, X itself being an element of B. It is assumed throughout this paper that B is separable; that is B is the Borel extension of a denumerable family of sets. The transition law of the process is given by a function $F(x,E) = P^1(x,E)$, this function being interpreted as the conditional probability that $x_{n+1} \in E$, given $x_n = x$. The n-step transition probability is designated by $P^n(x,E)$. When conditional probabilities are used below, it will usually be understood that they are the ones uniquely determined by the transition probabilities. The sets in B will sometimes be called "measurable sets".

Throughout this paper a "measure" will mean a countably additive set function, defined on the measurable sets, non-negative, and not identically 0. (The words "countably additive" will sometimes be repeated for emphasis.) A "probability measure" or "probability distribution" will be a measure of total mass 1. Notice that we do not require measures to be finite. A "sigma-finite" measure is a measure such that X is the union of a denumerable number of sets, each of which has a finite measure.

Various conditions are known which imply the existence of a probability measure $Q(E)$ which is a stationary distribution for the x_n-process; that is, Q satisfies, for every measurable E,

$$(1.1) \qquad Q(E) = \int_X Q(dx)P(x,E).$$

If x_0 has this distribution, so has x_n for every n. Two sets of conditions for the existence of such a probability measure were given by Doeblin. One set is discussed in Doob [7, pp. 190 ff]. A more general set is given in [6].

There are many situations where there is no probability measure satisfying (1.1), but where a solution can be found if $Q(X) = \infty$ is allowed. The simplest example is the random walk where x_n takes integer values, and can increase or decrease by 1, with probabilities 1/2 each, at each step. In this case, a solution to (1.1) can be obtained by assigning to any set of integers a Q-measure equal to the number of integers in the set. All integers are "equally probable."

In this paper a solution of (1.1) will always mean a sigma-finite measure Q which satisfies (1.1) for every measurable set E. The principal result, contined in Theorem 1, is that a certain type of recurrence condition on the x_n-process insures the existence of a solution of (1.1) which is unique, and which assigns positive measure to certain "recurrent" sets.

It is helpful to consider some known results for the case where X is denumerable, with states designated by letters $i,j,k,s,\ldots$, which are taken to be integers $0,1,\ldots$. Let

$P_{ij} = P_{ij}{}^1$ be the one-step transition probability from i to
j, and let $P_{ij}{}^n$ be the n-step transition probability. Assume
that all states communicate; i.e., for each i,j there is an n
such that $P_{ij}{}^n$ is positive. Assume also that the chain is
"recurrent"; that is, with probability 1, every state is visited
infinitely often. (If this is true for any starting point,
it is true for all starting points.) See Feller [9, Ch. 15]
for discussion. Under these conditions, it follows from a
result of Doeblin [5] that there is a set of positive finite
numbers q_j, $j = 0,.,...$, such that

$$(1.2) \qquad \lim_{N \to \infty} \sum_{n=1}^{N} P_{ij}{}^n \Big/ \sum_{n=1}^{N} P_{ks}{}^n = q_j/q_s$$

Notice that the limit in (1.2) is independent of the starting
positions i and k. In case the mean time for return of the
state back to the starting state is finite (under the above
hypothesis of "communication" this will be true for either all
or no states) the quantities q_i are a set of stationary pro-
babilities for the Markov chain, provided they are scaled so
their sum is 1. If the mean times of return are infinite
(still considering the recurrent case) the q_j are not pro-
babilities since their sum is infinite. However, it was shown
by Derman [3] that even in this case, the q_j satisfy the equations
of stationarity (1.3), which are of course satisfied also when
the q_i are probabilities:

$$(1.3) \qquad q_1 = \sum_{j=0}^{\infty} q_j P_{j1}, \quad 1 = 0,1,2,\ldots.$$

Another proof of (1.3) in the non-probability case was given by Chung [2] and furnishes an idea which is very helpful in the present proof.

There are several applications of infinite stationary measures Q. In the case where X is denumerable, one can consider infinitely many particles moving through the states in X, each one moving independently of the others according to the law of the Markov chain. In such cases Derman showed [4] that if initially the number of particles in state j has a Poisson distribution with mean q_j, $j = 0,1,\ldots$, this same distribution will be preserved as time goes on, and related results are also given. A similar treatment for certain types of continuous-parameter processes has been given by Doob, [7, p. 404].

Another application was discussed by Harris and Robbins, [14], where the application of an ergodic theorem due to Hopf, [15] shows that if f and g are real-valued functions such that the right side of (1.4) below is defined, then assuming a suitable recurrence condition holds, it will be true that for almost all starting states x_0

$$(1.4) \qquad \text{Prob}\left[\lim_{n\to\infty} \frac{f(x_0) + \ldots + f(x_n)}{g(x_0) + \ldots + g(x_n)} = \frac{\int f(x)Q(dx)}{\int g(x)Q(dx)} \right] = 1.$$

The statement "almost all" is with respect to Q-measure. In

some cases, such as the one discussed in the present paper, the exceptional set is empty.

The analogue of (1.4) for discrete Markov chains was proved by Chung without the use of ergodic theory.

A question which will be mentioned only briefly in this paper is whether the analogue of (1.2) holds for the more general state space X introduced above. That is, is it true that if $0 < Q(F) < \infty$ then

$$(1.5) \qquad \sum_{n=1}^{N} P^n(x,E) / \sum_{n=1}^{N} P^n(y,F) \longrightarrow Q(E)/Q(F)$$

holds for all x and y in X and all E in B? Conditions are known under which this is true; however, they seem to be less general than those under which Theorem 1 holds.

Finally it may be noted that even in case $Q(X)$ is finite, the result of the present paper applies to certain cases where one of the results of Doeblin mentioned above is not applicable (however, the conclusions are weaker than they would be if Doeblin's result could be applied) and where Doeblin's other result seems difficult to verify.

2. Recurrence Condition; the "Process on A"

We consider Markov processes as defined in the first paragraph of the introduction. Condition C: a countably addive sigma-finite measure $m(E)$ is defined on sets E of B; $m(E) > 0$ implies

$$\text{Prob } (x_n \in E \text{ infinitely often } | \ x_0) = 1$$

for all starting points x_0 in X.[1]

Now let A be a measurable set with $m(A) > 0$. Let x_0 be any point in A. Then almost all sequences x_0, x_1,..., will have infinitely many elements in A. Let $y_0 = x_0, y_1, y_2, \dots,$ be the successive members of the sequence which belong to A. It can be verified that the y_n form a Markov process[2] and we can write down its transition function $P_A(x, E) = \text{Prob } (y_{n+1} \in E \ | \ y_n = x)$:

$$(2.1) \qquad P_A(x, E) = P(x, E) + \int_{X-A} P(x, dy) P(y, E) + \int_{X-A} \int_{X-A} P(x, dy) P(y, dz) \cdot$$
$$\cdot P(z, E) + \dots.$$

We shall refer to the y_n-process as the process on A, or the A-process.

---

[1] Conditional probabilities will usually be those defined uniquely by the transition function.

[2] This intuitively obvious statement perhaps needs proof. It can readily be shown that the joint distribution of $y_1, y_2, \dots, y_n$ corresponds to what would be obtained by iteration of P_A, for all n.

3. Motivation for Proof of Theorem 1

It is evident from (1.2) that in the discrete case the quantities q_j are proportional to the expected number of times the state is j, over a long time interval. It follows from a proof of (1.2) by Chung [2] that the ratio q_j/q_1 is equal to the expected number of visits to j between two visits to 1, where $1 \neq j$. Hence we might think of taking Q(E) proportional to the expected number of visits to the set E by a "particle" which starts at some point x, before returning to x. The difficulty is that in general the particle does not have probability 1 of returning to the state x. It then seems natural to take some reference set A, with $m(A) > 0$, let the particle have some initial probability distribution on A, and take Q(E) to be the expected number of visits to E before returning to A. In general, there would be difficulties with this procedure, but the situation becomes simple if there is a stationary probability distribution for the A—process; if the particle has this distribution initially, it will have the same distribution each time it returns. This type of argument suggests the reason for Lemma 1 below. See also Levy, [18], where the A—process is used for discrete X, and Harris, [12]. The procedure used by Halmos, [11] in going from "bounded" to "sigma-bounded" spaces is related.

4. Proof of Existence of Stationary Measure

Theorem 1. <u>Assume the Markov process satisfies the conditions</u> <u>of paragraph 1 of the Introduction, and that Condition C</u> <u>holds. Then (1.1) has a solution Q which is unique up to a</u> <u>constant positive multiplier, and is stronger</u>[12] <u>than m.</u> (The precise meaning of "solution" is given in the Introduction).

A series of lemmas will be needed for the proof. The assumptions of Theorem 1 will be made for each of them, although Lemma 1 actually is true under wider conditions.

Lemma 1. <u>Let A be a measurable set with</u> $0 < m(A) < \infty$. <u>Suppose the "process on A" has a stationary probability measure</u> Q_A <u>satisfying</u>

$$(4.1) \qquad Q_A(E) = \int_A Q_A(dx) P_A(x, E)$$

<u>for every measurable</u> E <u>in A. Then (1.1) has a solution Q</u> <u>which assigns the same value as</u> Q_A <u>to subsets of A.</u>

Proof of Lemma 1. The function $P_A(x, E)$, which we defined before as the transition function for the A—process, can be defined by means of the right side of (2.1) for every x in X, and for all measurable E, whether or not E is a subset of A. The function so defined is clearly, for each x, a countably additive measure. An argument similar to the one we shall use below to show that Q is sigma-finite shows that P_A is a sigma-finite measure. Note that if E is a subset of X—A, then $P_A(x, E)$ is the expected number of visits to the set E

----

[12]That is, $m(E) > 0$ implies $Q(E) > 0$.

before a visit to A, if the starting point is x. (If $x \in A$, the initial position is not counted as a "visit" to A.) Now let A' denote the set X—A. From the definition of P_A it is clear that the following functional equation is satisfied:

$$(4.2) \qquad P_A(x,E) = P(x,E) + \int_{A'} P_A(x,dy)P(y,E).$$

Motivated by the considerations indicated in Section 3 above, we now define, for every measurable E in X,

$$(4.3) \qquad Q(E) = \int_A Q_A(dx)P_A(x,E).$$

Notice that for $E \subset A$, $Q(E) = Q_A(E)$. Using (4.1), (4.2), and (4.3), we have

$$(4.4) \qquad \int_X Q(dx)P(x,E) = \int_A Q_A(dx)P(x,E) + \int_{A'} [\int_A Q_A(dx)P_A(x,dy)]P(y,E) =$$

$$\int_A Q_A(dx)[P(x,E) + \int_{A'} P_A(x,dy)P(y,E)] = \int_A Q_A(dx)P_A(x,E) = Q(E).$$

Hence Q satisfies (1.1), and is clearly a countably additive measure.

To show that Q is sigma-finite, we note that $Q_A(A) = Q(A) = 1$, so it is sufficient to show that A' is the union of a denumerable number of sets, each with finite Q—measure. Define $S_{i,j}$ for every pair of positive integers i and j by

$$(4.5) \qquad S_{i,j} = \left\{ x : x \in A', P^i(x,A) > 1/j \right\}, \quad i,j = 1,2,\ldots$$

From Condition C it follows that the union of the S_{1j} contains
A'. Since $S_{1j} \subset A'$, $Q(S_{1j})$ is just the expected number of visits
to S_{1j} between visits to A, if the "process on A" has the
stationary probability distribution Q_A. It is therefore
clear that $Q(S_{1j})$ is finite.

It remains only to show that Q is stronger than m. Since
Q_A is by assumption stronger than m on A, and Q coincides with
Q_A on A, it will suffice to show that Q is stronger than m
on A'. Hence let E be a measurable subset of A', with $Q(E) = 0$.
Suppose the x_n-process, $n \geq 0$, has the initial distribution
Q_A on A, x_0 being assigned probability 0 of being in A'.
Since $Q(E)$ is the expected number of visits to E between visits
to A, $Q(E) = 0$ evidently implies that with probability 1,
x_n does not belong to E for any n. This contradicts Condition
C, which implies that for any starting point x, there is pro-
bability 1 of visiting E infinitely often.

This completes the proof of Lemma 1.

We next consider the absolutely continuous parts (with
respect to the measure m) of $P(x,E)$ and its iterates. For
each positive integer n, let

$$(4.6) \qquad P^n(x,E) = \int_E f^n(x,y) m(dy) + P_0^n(x,E),$$

where $f^n(x,y)$ is thus the density, and P_0^n is the singular
component. Our original condition that the field B should be
seperable implies that the representation (4.6) is possible
with f^n a function which is measurable in the pair (x,y),

measurability being defined with respect to the product space (X,X). See Doob, [7, pp. 196 and 616]. We shall henceforth assume that the densities satisfy this measurability condition.

Lemma 2. Let r be any real number strictly between 0 and 1. There exist a measurable set A, a positive number s, and a positive integer k, such that $0 < m(A) < \infty$, and for every $x \in A$

$$(2.4) \qquad m \left\{ y: y \in A, \; f^1(x,y) + \ldots + f^k(x,y) > s \right\} > r \, m(A).$$

Proof of Lemma 2. Since each of the measures P_0^n, $n > 0$, is singular, we can find, for each x, a measurable set $S(x)$ with $m(S(x)) = 0$, such that

$$(4.7) \qquad P_0^n(X - S(x)) = 0, \; n = 1,2,\ldots.$$

For each x, let $T(x)$ be the (measurable) x-set defined by

$$(4.8) \qquad T(x) = \left\{ y: f^n(x,y) = 0, \; n = 1,2,\ldots. \right\}.$$

Then if $x_0 = x$, the probability is 1 that there is no n such that x_n belongs to $T(x) - T(x)S(x)$. Hence from Condition C, $m(T(x)) = 0$.

Now let A_1 be any measurable set such that $0 < m(A_1) < \infty$. For each $x \in A_1$, define the measurable set $A_{1i} = A_{1i}(x)$ for $i = 1,2,\ldots,$ by

$$(4.10) \qquad A_{1i}(x) = \left\{ y: y \in A_1, \; f^1(x,y) + \ldots + f^i(x,y) > 1/i \right\}.$$

The remark above that $m(T(x)) = 0$ shows that

$$(4.11) \qquad m(A_1 - U_1 A_{11}) = 0.$$

Hence for each $x \in A_1$ there is a smallest positive integer $i = i(x)$ such that

$$(4.12) \qquad m(A_{11}) > (\frac{1}{2} + \frac{1}{2}r) \, m(A_1).$$

Now define

$$(4.13) \qquad K_j = \{ x : x \in A_1, \; i(x) = j \}, \; j = 1, 2, \ldots$$

Then each x in A_1 belongs to some K_j. Hence we can find an integer p such that

$$(4.14) \qquad m(K_1 + \ldots + K_p) > 2m(A_1)/(4 - r).$$

Now define $A = K_1 + \ldots + K_j$. Then for every x in A the relation

$$(4.15) \qquad f^1(x,y) + \ldots + f^p(x,y) > 1/p$$

holds for all y in A_1 with the possible exception of a y-set
: m-measure not exceeding $\frac{1}{2}(1 - r) \, m(A_1)$. But $m(A_1) <$
$\frac{1}{2}(4 - r) m(A)$. Hence for every $x \in A$, (4.15) must hold for all y
in A with the possible exception of a y-set of m-measure not
exceeding $[(1 - r)(4 - r)/4] \, m(A) < (1 - r) \, m(A)$. Hence
(4.7) is true with $\kappa = p$ and $s = 1/p$. This completes the proof
of Lemma 2.

Lemma 3. Let A be the set of Lemma 2, corresponding to a
particular value of r. Suppose that the transition functions

for the "process on A" have the decomposition (satisfying
the measurability condition mentioned in connection with (4.6))

$$(4.16) \qquad P_A^n(x,E) = \int_E f_A^n(x,y)\, m(dy) + P_{A0}^n(x,E), n=1,2,\ldots, x \in A,\ E \subset A.$$

Then (4.7) still holds, with the same k and s, if $f^1 + \ldots + f^k$
is replaced by $f_A^1 + \ldots + f_A^k$

    Proof of Lemma 3. We observe that, for $E \subset A$, $P^1(x,E) +$
$\ldots + P^n(x,E)$ is the expected number of visits to E in n
steps, while $P_A^1(x,E) + \ldots + P_A^n(x,E)$ is the expected number
of visits to E in the first n visits to A. Hence clearly for
every $x \in A$ and $E \subset A$

$$(4.17) \qquad P_A^1 + \ldots + P_A^n \geq P^1 + \ldots + P^n, \quad n = 1,2,\ldots$$

Let A_2 be a subset of A with $m(A_2) = m(A)$, such that

$$(4.18) \qquad P_{A0}^n(x,A_2) = 0, \quad n = 1,2,\ldots$$

Then for $E \subset A_2$ (A_2 may depend on x)

$$(4.19) \qquad \int_E (f_A^1(x,y) + \ldots + f_A^n(x,y)) m(dy) \geq \int_E (f^1(x,y) + \ldots + f^n(x,y)) m(dy).$$

Hence for each x, the relation $f_A^1 + \ldots + f_A^n \geq f^1 + \ldots + f^n$
must hold for almost all y in A (m-measure.) Lemma 3 follows
immediately from this fact.

Lemma 4. Let $R(x,E)$ be the Markov transition function defined by

$$(4.20) \qquad R(x,E) = [P_A^1(x,E) + \ldots + P_A^k(x,E)]/k, x \in A,\ E \subset A,$$

where A, $P_A(x,E)$, and R are the quantities defined in Lemmas 2 and 3 for some $r > 1/\sqrt{2}$. Let R^n, $n = 1,2,\ldots$, be the iterates of R. Then there is a stationary probability distribution $Q_A(E)$ for P, satisfying, for all measurable $E \subset A$,

$$(4.21) \qquad Q_A(E) = \int_A Q_A(dx) \, R(x,E).$$

Moreover there is a number t, $0 < t < 1$, such that if $\phi(E)$ is any probability measure on A,

$$(4.22) \qquad \left| Q_A(E) - \int_A \phi(dx) R^n(x,E) \right| \leq t^{n-1}, \quad n = 1,2,\ldots.$$

The proof of Lemma 4 follows from results of Doeblin; see Doob, [7, pp. 190 ff.] However, the transition function R satisfies such a strong positivity condition that a very simple proof of Lemma 4 can be given. Such a proof is given in the Appendix to the present paper.[14]

Lemma 5. The distribution $Q_A(E)$ of Lemma 4 is stationary for $P_A(x,E)$; that is, (4.1) is satisfied.

Proof of Lemma 5. Let us define operators U and V, on any probability measure ϕ defined on A, by

$$(4.23) \qquad U\phi(E) = \int_A \phi(dx) P_A(x,E); \quad V\phi(E) = \int_A \phi(dx) \, R(x,E).$$

---

[14] If Doeblin's result is to be used directly, it is sufficient to take r greater than $\frac{1}{2}$ in Lemma 4.

$U\phi$ and $V\phi$ are clearly themselves probability measures. The nth power of the operator U (or V) is obtained by replacing P_A by $P_A{}^n$ (or R by R^n). Note also that because of the definition of R (see (4.20)), we have

$$(4.24) \qquad V = (U + U^2 + \ldots + U^k)/k$$

From (4.24) we see that U and V commute: $U^n V^m = V^m U^n$ for any positive integers m and n. Now (4.21), in the present language, becomes

$$(4.25) \qquad Q_A = VQ_A,$$

and (4.22) states that for any probability measure ϕ defined on A

$$(4.26) \qquad |V^n\phi(E) - Q_A(E)| \leq t^{n-1}, \quad n = 1,2,\ldots.$$

Using (4.25), (4.26), and commutativity, we have (the measures in (4.27) are evaluated for some fixed measurable set E)

$$(4.27) \qquad UQ_A = UV^n Q_A = V^n(UQ_A) = Q_A + \mathcal{E}_n$$

where $\mathcal{E}_n$ is a quantity which is less than t^{n-1} in magnitude. This can only be true if $UQ_A = Q_A$, which is the desired result.

In order to apply Lemma 1, we need to know that Q_A is stronger than m on A. Suppose that E is a measurable subset of A for which $Q_A(E) = 0$. Consider the "process on A" $y_0, y_1, \ldots,$ where y_0 has the distribution Q_A. Because of the stationarity

of the y_n-process, we have

(4.28)   Prob $(y_n \in E)$ = $Q_A(E)$ = 0, n = 1,2,....

Because of Condition C, (4.28) implies $m(E)$ = 0.

We can now apply Lemma 1. There is then a solution Q
of (1.1); we recall that a solution is a countably additive,
sigma-finite measure; Lemma 1 shows it to be stronger than m.

It remains to prove uniqueness. We shall show that, except
for a multiplicative constant, there is only one sigma-finite
measure which satisfies (1.1), even among measures which are
not necessarily stronger than m. It is convenient to defer
the proof of uniqueness to the next section.

5. Uniqueness and Related Topics

Throughout Section 5 we shall still assume that the conditions of the first paragraph of the Introduction hold, and that Condition C is satisfied. Furthermore, it will be assumed that a solution $Q(E)$ of (1.1) has been constructed by use of some fixed set A, as in Lemmas 1-5; this solution has been shown to be sigma-finite and stronger than m, and we shall reserve the notation Q for it throughout the present section.

Lemma 6. _Let_ D _be a measurable set with_ $0 < m(D)$, $Q(D) < \infty$. _Then_ $Q(E)/Q(D)$ _is a stationary probability measure for the_ "process on D", _satisfying_ (P_D _is the transition function for the D-process_)

$$(5.1) \qquad Q(E)/Q(D) = \int_D [Q(dx)/Q(D)] P_D(x,E), \quad E \subset D.$$

It is the only stationary probability measure for this process.

We mention explicitly that Q was constructed with reference to a fixed set A, not with reference to D.

Proof of Lemma 6. For simplicity, take $Q(D) = 1$. Let $D' = X - D$. Then, taking (1.1) as our starting point we obtain

$$(5.2) \qquad Q(E) = \int_X Q(dx) P(x,E) = \int_D + \int_{D'}, \quad E \subset D.$$

The second term on the right side of (5.2) can be written as

$$(5.3) \qquad \int_{D'} Q(dy) P(y,E) = \int_{D'} [\int_X Q(dx) P(x,dy)] P(y,E) =$$

$$\int_X Q(dx) \int_{D'} P(x,dy) P(y,E) = \int_D \int_{D'} + \int_{D'} \int_{D'}$$

Hence we have

$$(5.4) \qquad Q(E) = \int_D Q(dx) \left[P(x,E) + \int_{D'} P(x,dy) P(y,E) \right] + \int_{D'} \int_{D'},$$

or

$$(5.5) \qquad Q(E) \geq \int_D Q(dx) \left[P(x,E) + \int_{D'} P(x,dy) P(y,E) \right].$$

This process can be continued, replacing $Q(dx)$ in the last term in (5.4) by

$$(5.6) \qquad \int_X Q(dy) P(y,dx).$$

Hence we obtain

$$(5.7) \qquad Q(E) \geq \int_D Q(dx) \left[P(x,E) + \int_{D'} P(x,dy) P(y,E) + \ldots \right]$$

$$= \int_D Q(dx) \, P_D(x,E), \quad E \subset D.$$

Now $Q(E)$ is a probability measure on D (we have taken $Q(D) = 1$). Since P_D is a Markov transition probability function, the right side of (5.7) is likewise a probability measure on D. Hence (5.7) can only be true if the equality (5.1) holds.

At this point we observe that the only property of Q of which we have made use is the fact that it is a solution of (1.1). The above argument will then go through for any measurable set D such that $0 < Q(D) < \infty$, and such that Prob $(x_n \in D$ inf. often $| x_0) = 1$ for all x_0.

Now consider a "process on D" with variables z_n, $n = 0, ., \ldots,$ where for each n, z_n has the stationary distribution Q on D. Suppose D is the union of two disjoint nonempty measurable sets $D_1 \cup D_2$. Since $m(D) > 0$, either D_1 or D_2 must have positive

m-measure. It follows that the two statements in (5.8) cannot simultaneously be true:

(5.8)   Prob $(z_n \in D_1, n=1,2,\ldots |z_0=z)=1$, all $z \in D_1$,

          Prob $(z_n \in D_2, n=1,2,\ldots |z_0=z)=1$, all $z \in D_2$.

The fact that two such statements cannot be made is stronger than mere ergodicity and in conjunction with the fact that Q is stronger than m, implies the following proposition. Let E be a measurable subset of D and let $g(z)$ be the characteristic function of E. Then for all (not just almost all) z in D we have

$$(5.9) \qquad \text{Prob}\left\{ \lim_{n \to \infty} \frac{g(z_0)+\ldots+g(z_{n-1})}{n} = Q(E) \;|z_0 = z \right\} = 1.$$

(See the discussion of a similar situation in Doob, [7 Theorem 6.2, p. 220]). Now any stationary probability distribution Q' defined on D for the D-process will correspond to an ergodic[15] stationary process z_n, where z_n has the distribution Q'. The ergodic theorem of Birkhoff then implies that (5.9) holds, with $Q(E)$ replaced by $Q'(E)$ for almost all (Q'-measure) z in D. (We cannot say all z at this point since a priori we do not know that Z' is stronger than m.) This is a contradiction unless Q' and Q are equal. Hence Q is unique among

───────────────

[15]Ergodicity follows as for the Q-process.

stationary probability measures for the D-process. This completes
the proof of Lemma 6.

We can get some additional information from (5.9).
We know that Q is stronger than m, and it may be strictly
stronger; that is, there may be sets E with $m(E) = 0$, $Q(E) > 0$.
It can be shown from (5.9) that $Q(E) > 0$ implies

(5.10)   Prob $(x_n \in E \text{ i.o.} | x_0) = 1$, all $x_0 \in X$.

We can now conclude the uniqueness proof for the solution
of (1.1). Suppose Q is our solution and that Q_1 is any different
solution. Then there must be a set D such that

(5.11) $0 < m(D), Q(D), Q_1(D) < \infty$,

(for convenience let us take $Q(D) = Q_1(D) = 1$) and such that for
some subset E of D, we have $Q(E) \neq Q_1(E)$. From the discussion
above it follows that both Q and Q_1 must be stationary probability
distributions on D for the D-process; the fact that Q is such
a distribution was proved; the fact that Q_1 is such a distri-
bution follows from (5.10) and the remark made in the course
of the proof of Lemma 6, to the effect that any solution of
(1.1) provides a stationary distribution for the D-process
under the indicated conditions. Since from Lemma 6 the station-
ary distribution is unique, we have a contradiction.

This completes the proof of Theorem 1.

6. Remarks

Theorem 1 is sometimes helpful in the finite case, $Q(X) < \infty$. It is useful to know in advance whether the finite or infinite case prevails. The following criterion will be stated without proof. More general ones could clearly be found.

Suppose the conditions of Theorem 1 are satisfied, and suppose we can find a set A, $0 < m(A) < \infty$, and a positive number d such that the density $f(x,y)$ satisfies

$$(6.1) \qquad f(x,y) > d, \quad x \in A, \; y \in A.$$

Let ρ be the recurrence time to A; that is, ρ is the smallest positive integer such that $x_\rho \in A$. Then if $E(\rho \,|\, x_0 = x)$ is bounded independently of x for all $x \in A$, we have the finite case.

Although some conditions have been found under which (1.5) is true, (see [13], the author has not been able to prove this under the conditions of Theorem 1 alone.

It does not seem easy to apply the general ergodic theorems of Halmos [10] and [11], Dowker [8], and Hurewicz, [17] to the present problem. See also Hopf, [16]. However, it would be interesting to explore possible connections. As mentioned in the introduction, the idea of extending a finite measure on a subset to an infinite measure on the whole space is somewhat related . a procedure used by Halmos, [.1].

Some recent results of Edward Nelson, in a thesis at

the University of Chicago, may have connections with the

present problem. The author has as yet been unable to see

this work.

APPENDIX

Consider $R(x,E)$ defined by (4.20), and let $g(x,y)$ be its density with respect to m. Then there is a positive number, say b, such that for all $x \in A$,

$$(A.1) \qquad m\{y : g(x,y) > b\} > rm(A) > m(A)/\sqrt{2},$$

where r is the number of Lemma 2, taken here to be $> 1/\sqrt{2}$. Now define

$$(A.2) \qquad M_n(E) = \sup_{x,y} (R^n(x,E) - R^n(y,E))$$

$$W_n = \left\{(u,v) : R^n(u,E) - R^n(v,E) > 0\right\}$$

where all points in the definitions are in A. Furthermore, let A_2 denote the product space (A,A), and let m_2 denote the product measure in this space constructed from m. Let $G(x,y)$ be the subset of A_2 defined by

$$(A.3) \qquad G(x,y) = \left\{(u,v) : g(x,u)g(y,v) > b^2\right\}$$

Then $m_2(G) > r^2 m_2(A_2) = r^2(m(A))^2$. Now for each pair (u,v) where the expression defining W_n is positive, there is a corresponding pair (v,u) where the expression is negative, and since the measure is product measure, these two mutually-exclusive sets have equal m_2-measures. Hence $m_2(A_2 - W_n) \geq \frac{1}{2} m_2(A_2)$. Hence

$$(A.4) \qquad m_2 \lfloor G(A_2 - W_n) \rfloor > (r^2 - \frac{1}{2})m_2(A_2).$$

With these preliminaries we can write

$$
(A.5) \qquad R^{n+1}(x,E) - R^{n+1}(y,E) = \int_A R(x,du)R^n(u,E) -
$$

$$
\int_A R(y,dv)R^n(v,E) = \iint_{A \; A} R(x,du)R(y,dv)(R^n(u,E)-R^n(v,E)) \le
$$

$$
\int \int_{W_n} R(x,du)R(y,dv)(R^n(u,E)-R^n(v,E)) \le
$$

$$
M_n(E) \iint_{W_n} R(x,du)R(y,dv).
$$

Hence

$$
(A.6) \qquad M_{n+1}(E) \le M_n(E) \sup_{x,y} \iint_{W_n} R(x,du)R(y,dv)
$$

Now

$$
(A.7) \qquad \iint_{A_2-W_n} R(x,du)R(y,dv) \ge
$$

$$
\iint_{A_2-W_n} g(x,u)g(y,v)m(du)m(dv) \ge
$$

$$
b^2(r^2-\tfrac{1}{r}) \; m_2(A_2) = c > 0.
$$

Note also that the double integral on the right side of (A.6)
would be 1 if it were taken over the whole space A_2. Hence, from
(A.7) we have

$$
(A.8) \qquad \sup_{x,y} \int \int_{W_n} R(x,du)R(y,dv) \le 1 - c, \; 0 < 1 - c < 1.
$$

It follows that

$$
(A.9) \qquad M_n(E) \le (1-c)^{n-1}, \; n = 1,2,\dots.
$$

The same type of argument shows that the absolute value of the infimum on x and y of the quantity $R^n(x,E) - R^n(y,E)$ satisfies the inequality in (A.9). The desired result follows readily from this.

REFERENCES

1. Chung, K. L., "Contributions to the Theory of Markov Chains I", Journal of Research of the National Bureau of Standards, Vol. 50 (1953), pp. 203-208.

2. _____, "Contributions to the Theory of Markov Chains II", Transactions of the Am. Math. Soc., Vol. 76 (1954), pp. 397-419.

3. Derman, C., "A Solution to a Set of Fundamental Equations in Markov Chains", Proc. Am. Math. Soc., Vol. 5 (1954), pp. 332-334.

4. _____, "Some Contributions to the Theory of Markov Chains", Trans. Am. Math. Soc., to appear.

5. Doeblin, W., "Sur Deux Problems de M. Kolmogoroff", Bull. Soc. Math. France, Vol. 52 (1938), pp. 210-220.

6. _____, "Elements d'une Theorie Générale Des Chaînes Simples Constantes De Markoff", Ann. Ecole Norm. Sup. (3), LVII, Fasc. 2, (1940), pp. 61-111.

7. Doob, J. L., Stochastic Processes, New York, Wiley, 1953.

8. Dowker, Y. N., "Finite and Sigma-Finite Invariant Measures", Ann. Math., Vol. 54 (1951), 595-608.

9. Feller, W., An Introduction to Probability Theory and its Applications, New York, Wiley, 1950.

10. Halmos, P. R., "An Ergodic Theorem", Proc. Nat. Acad. Sci. USA, Vol. 32 (1946), pp. 156-161.

11. _____, "Invariant Measures", Ann. Math., Vol. 48 (1947), pp. 735-754.

12. Harris, T. E., "First Passage and Recurrence Distributions", Trans. Am. Math. Soc., Vol 73 (1952), pp. 471-486.

13. _____, "Recurrent Markov Processes", Abstract, Ann. Math. Stat., Vol. 26 (1955), p. 152.

14. _____, and Robbins, H. E., "Ergodic Theory of Markov chains admitting an infinite invariant measure", Proc. Nat. Acad. Sci. USA, Vol. 39 (1953), pp. 860-864.

15. Hopf, E., Ergodentheorie (reprinted), New York, Chelsea, 1948, pp. 46-53.

16. _____, "Theory of Measure and Invariant Integrals", Trans. Am. Math. Soc., Vol. 34 (1932), pp. 373-393.

17. Hurewicz, W., "Ergodic Theorem Without Invariant Measure", Ann. Math., Vol. 45 (1944), pp. 192-206.

18. Levy, P., "Systemes Markoviens et Stationnaires; cas Denombrable", Ann. Ecole Norm. (3), Vol. 68 (1952), pp. 327-381.